

基于加权双域稀疏分解模型的编码器信号特征提取 及其行星齿轮箱故障诊断应用^{*}

张建峰¹, 丁传仓², 江星星², 杜贵府², 李舜酩¹

(1. 南通理工学院汽车工程学院 南通 226002; 2. 苏州大学轨道交通学院 苏州 215131)

摘要:齿轮箱作为机械设备中的关键传动部件,常年在严苛环境和重载条件下工作,易发生损伤。近年来,内置编码器信号因其成本低廉、获取便捷等优势,逐渐成为机械设备健康状态监测的重要手段。然而,由于编码器信号中的故障特征微弱,且易与多种干扰成分交叠,导致故障冲击特征难以准确辨识与提取。针对上述挑战,提出了一种基于加权双域稀疏分解模型的编码器信号特征提取及故障诊断方法。在加权双域稀疏分解模型中,分析了故障冲击特征与干扰成分在时域和频域中的形态特性差异,随后通过引入权重系数、非凸惩罚函数及周期二进制向量,构建了分别约束时域中故障冲击特征的周期性组稀疏特性与频域中谐波干扰成分的频域稀疏特性的非凸正则项,实现了信号中目标特征的精准匹配与刻画。此外,采用交替方向乘子法及受控极小化方法,推导了加权双域稀疏分解模型的迭代求解算法。最后,通过仿真数据与行星齿轮箱实验台数据验证了加权双域稀疏分解模型的有效性。结果表明,相较于两种对比方法——调Q因子小波变换分解方法和最大相关峭度解卷积方法,该方法在故障冲击特征提取精度方面具有显著优势,展现出良好的工程应用潜力。

关键词:加权双域稀疏分解模型;编码器信号;故障诊断;特征提取;齿轮箱

中图分类号: TH165+.3 TH133.33 文献标识码: A 国家标准学科分类代码: 460.99

Weighted bi-domain sparse decomposition model for feature extraction of encoder signals and its application in planetary gearbox fault diagnosis

Zhang Jianfeng¹, Ding Chuancang², Jiang Xingxing², Du Guifu², Li Shunming¹

(1. School of Automotive Engineering, Nantong Institute of Technology, Nantong 226002, China;
2. School of Rail Transportation, Soochow University, Suzhou 215131, China)

Abstract:As a key transmission component in mechanical systems, gearboxes often work under harsh environmental and heavy-load conditions, making them prone to damage. In recent years, built-in encoder signals have emerged as an effective method for health monitoring of mechanical equipment due to their low cost and ease of acquisition. However, the fault-related features within encoder signals are typically weak and often masked by various interference components, which poses a significant challenge for accurate identification and extraction of fault-induced impulses. To address this issue, this article proposes a fault feature extraction and diagnosis method for encoder signals based on a weighted bi-domain sparse decomposition model. In the proposed model, the morphological differences between fault impulses and interference components are analyzed in both the time and frequency domains. By introducing weighted coefficients, periodic binary vectors, and non-convex penalty functions, the model constructs two dedicated non-convex regularization terms that enforce periodic group sparsity in the time domain for fault impulses and spectral sparsity in the frequency domain for harmonic interferences, thereby achieving accurate separation and representation of target features. Furthermore, an iterative solving algorithm for the weighted bi-domain sparse decomposition model is developed by integrating the alternating direction method of multipliers and the majorization-minimization strategy to enhance computational efficiency and convergence stability. Finally, the effectiveness of the proposed weighted bi-domain sparse decomposition model is validated using both simulated data and experimental data

from a planetary gearbox test bench. The results indicate that, compared with two existing methods, namely the tunable Q wavelet transform-based decomposition method and the maximum correlated kurtosis deconvolution method, the proposed method shows a significant advantage in fault impulse feature extraction accuracy, demonstrating strong potential for practical engineering applications.

Keywords: weighted bi-domain sparse decomposition model; encoder signals; fault diagnosis; feature extraction; gearbox

0 引言

齿轮箱作为机械系统中实现动力传递与变速的关键部件,广泛应用于风电、冶金、矿山、轨道交通及航空航天等重大装备中,其运行状态直接关系到整机系统的性能、安全性和使用寿命。一旦齿轮箱发生故障,轻则降低传动效率、加剧机械磨损,重则可能导致设备停机甚至引发灾难性事故。因此,开展高效且可靠的齿轮箱故障诊断技术研究,对于实现设备状态监测、预知维护与故障预警具有重要的现实意义^[1]。

目前,基于振动信号分析的故障诊断方法在齿轮箱监测领域已被广泛采用^[2-4]。该类方法通常通过安装加速度传感器采集机械振动信号,利用信号处理手段提取故障特征,取得了良好的诊断效果。然而,在实际工程应用中,基于振动信号分析的故障诊断方法仍面临诸多挑战。一方面,振动信号极易受到多源干扰的影响,尤其在低转速与强背景噪声工况下,故障冲击信号能量微弱、信噪比较低,关键特征往往被噪声淹没;另一方面,传感器的布设受限于机械结构与工作环境,难以对设备内部关键部位进行有效监测^[5-6]。因此,研究一种无需额外传感器、对微弱故障高度敏感的替代监测手段具有重要的理论意义和工程价值。

近年来,随着控制精度要求的不断提升,大量旋转机械设备集成了高精度旋转编码器,用于获取角度、转速等运动学信息。研究表明,编码器信号中蕴含丰富的动力学响应特征,当机械部件在运行过程中发生异常时,将引起编码器信号产生相应变化。因此,若能有效提取编码器信号中的异常信息,即可实现对机械部件健康状态的识别^[7]。目前,基于编码器信号的故障诊断方法展现出良好的应用前景,逐渐成为齿轮箱状态监测研究中的热点方向。

为便于机械故障的识别,通常需将记录旋转轴角位移的原始编码器信号转换为更具诊断价值的运动学变量,如瞬时角速度(instantaneous angular speed,IAS)与瞬时角加速度(instantaneous angular acceleration,IAA)等^[8]。机械故障的存在将改变设备的动态行为,从而导致IAS、IAA发生周期性波动。例如,刘振等^[9]基于动力学建模分析了齿轮裂纹损伤对行星齿轮箱编码器信号响应特性的影响,并利用编码器信号开展齿轮箱健康状态监测。Li等^[10]提出结合经验模态分解与自相关局部倒

谱的编码器信号分析方法,成功实现多级齿轮箱的故障检测。Zhao等^[11]设计了一种集成于滚动轴承的摩擦电传感器,通过分析IAS信号的周期波动实现轴承故障诊断。这些研究表明,编码器信号在机械故障识别方面具有良好的可行性和潜力。

然而,编码器信号在实际应用中不可避免地受到各种干扰成分影响,从而导致其导出的IAS、IAA等变量信号亦存在干扰问题,进而影响故障诊断的精度。为解决该问题,Ding等^[12]分析了IAS和IAA在差分域中所表现出的周期组稀疏结构特征,构建了周期性组稀疏差分模型,用于编码器信号的稀疏优化去噪,有效提取出周期性瞬态特征。Wang等^[13]利用自重采样技术将时域IAS映射至角度域,并结合基于低通滤波和稀疏建模的算法提取故障引起的IAS波动,实现了在变转速条件下的故障诊断。欧曙东等^[14]针对低转速旋转机械的故障诊断难题,构建了稀疏低秩分解模型以分离低速下的故障冲击,并引入快速主成分追踪算法进行求解,实现了对低转速条件下机械故障的识别。田田等^[15]针对行星减速器太阳轮故障检测,提出基于改进自适应多点最优最小熵解卷积的方法,通过谱负熵最大化原则自适应优化滤波器长度,实现了太阳轮故障特征的准确提取。张佳鑫等^[16]将角度同步平均与自适应最大相关熵解卷积相结合,解决了工业机器人中RV减速器低速非整周期旋转下的针齿故障识别问题。Zuo等^[17]提出平稳特征模态分解方法,以最大二阶循环平稳性为最优目标精准锁定故障特征信息,结合熵准则对分解模态进行筛选,提升了齿轮箱特征提取的准确性。

尽管上述研究已在编码器信号分析方面取得了显著进展,但由于对信号结构内在特征的利用仍不充分,导致其在强干扰背景下的特征提取性能仍存在一定局限。具体而言,故障冲击信号在时间域中通常呈现为周期性组稀疏结构,而谐波干扰成分在频率域中则体现为典型的频域稀疏性。若能将上述不同形态特征引入到特征提取方法的设计中,构建同时兼顾时域与频域稀疏性的分解模型,有望实现不同成分的有效解耦,提高对弱故障信号的提取与识别能力。

在此背景下,本研究提出一种基于加权双域稀疏分解模型的编码器信号特征提取方法,旨在深入挖掘编码器信号中潜藏的微弱周期性故障冲击特征,实现对机械系统故障的准确诊断。在该模型中,通过引入权重系数、非凸惩罚函数及周期二进制向量,构建用于分别约束周

期性组稀疏特性与频域稀疏特性的正则项,从而实现对冲击成分与谐波干扰的靶向匹配与分离。同时,针对构建的加权双域稀疏分解模型,结合交替方向乘子法与受控极小化方法,推导出高效稳定的迭代求解算法。最后,通过仿真编码器信号与行星齿轮箱实验台数据对所提方法的性能进行验证,实验结果表明,提出方法在故障冲击特征提取的准确性优于现有对比方法,具有良好的工程应用前景。

1 理论基础

1.1 稀疏模型

稀疏表示核心思想是利用一个冗余但具有良好表达能力的字典 \mathbf{D} ,将原始信号表示为少量原子的线性组合。该方法基于信号在某些变换域中具有固有稀疏结构的假设,即大多数实际信号在某种域内仅需极少数的基向量即可精确重构。这一特性不仅能有效压缩冗余信息,更有助于突出信号的本征结构,适用于在强背景噪声干扰下提取微弱的冲击特征。

设观测信号为 $\mathbf{y} \in \mathbf{R}^N$,其可通过字典 $\mathbf{D} \in \mathbf{R}^{N \times K}$ 表示为:

$$\mathbf{y} = \mathbf{D}\mathbf{x} + \mathbf{n} \quad (1)$$

式中: $\mathbf{x} \in \mathbf{R}^K$ 是稀疏系数向量,要求其大部分元素为 0,仅少数非零元素参与信号的重构; \mathbf{n} 表示观测噪声或建模误差项。上述模型的目标是求解最稀疏的 \mathbf{x} ,即在满足数据拟合条件的前提下,实现信号的最简表示。

为了实现信号的最简表示,构建稀疏正则优化模型,即:

$$\min_{\mathbf{x}} \frac{1}{2} \|\mathbf{y} - \mathbf{D}\mathbf{x}\|_2^2 + \lambda \|\mathbf{x}\|_1 \quad (2)$$

式中: $\|\mathbf{x}\|_1$ 表示 l_1 范数,定义为 $\|\mathbf{x}\|_1 = \sum_i |x_i|$; $\lambda > 0$ 是正则参数,用于平衡重构误差与稀疏性惩罚项之间的比例。该模型在压缩感知、信号去噪及故障特征增强等领域得到了广泛应用。

1.2 非凸惩罚函数

在稀疏优化问题中,一般利用 l_1 范数约束稀疏性。然而, l_1 范数会使得较大值被低估,导致无法准确还原期望的特征成分。为克服这一局限,近年来研究者利用非凸惩罚函数替代传统的 l_1 范数以避免较大值被低估^[18],增强对关键稀疏成分的保留能力。

表 1 列出了 4 种典型的惩罚函数 $\phi(u; a)$,即凸惩罚函数 abs 和非凸惩罚函数 log、rat、atan,其中 u 为惩罚函数的输入变量,参数 a 控制惩罚函数的非凸程度。

图 1 显示了凸惩罚函数 abs 及 $a = 1$ 时的非凸惩罚函数 log、rat 和 atan。从图 1 可以看出,在参数 a 相同的

表 1 4 种典型惩罚函数

Table 1 Four typical penalty functions

惩罚函数	$\phi(u; a)$	$\phi'(u; a)/u$
abs	$ u $	$\frac{1}{ u }$
log	$\frac{1}{a} \log(1+a u)$	$\frac{1}{ u (1+a u)}$
rat	$\frac{ u }{1+a u /2}$	$\frac{1}{ u (1+a u)^2}$
atan	$\frac{2}{a\sqrt{3}} \left(\tan^{-1}\left(\frac{1+2a u }{\sqrt{3}}\right) - \frac{\pi}{6} \right)$	$\frac{1}{ u (1+a u +a^2 u ^2)^2}$

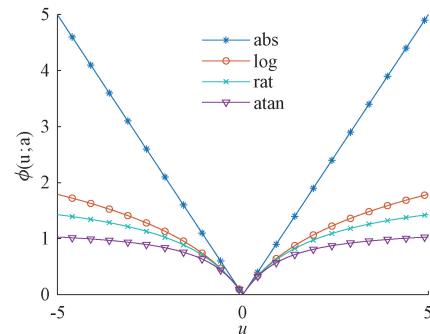


图 1 $a = 1$ 时的 4 种典型的惩罚函数

Fig. 1 Four typical penalty functions when $a = 1$

情况下,惩罚函数 atan 的非凸程度比其他惩罚函数更强,可以最大程度地促进稀疏性和保持较大值,因而更适合用于故障冲击特征的突出与恢复。

1.3 受控极小化方法

受控极小化方法是一种有效的稀疏优化求解策略,其核心思想是通过构建一系列相对简化的优化子问题,从而逐步逼近全局最优解^[19]。

对于一般形式的凸优化问题(式(3)),其最优解可以通过迭代方式获得。

$$\{\mathbf{u}^*\} = \arg \min_{\{\mathbf{u}\}} F(\mathbf{u}) \quad (3)$$

具体而言,令 $\mathbf{u}^{(j+1)}$ 表示第 $j+1$ 次迭代的估计值,则其迭代更新形式可表示为:

$$\{\mathbf{u}^{(j+1)}\} = \arg \min_{\{\mathbf{u}\}} Q(\mathbf{u}, \mathbf{u}^{(j)}) \quad (4)$$

式中: j 表示第 j 次迭代; $Q: \mathbf{R}^N \times \mathbf{R}^N \rightarrow \mathbf{R}^N$ 为目标函数 F 的受控优化算子,满足:

$$\begin{cases} Q(\mathbf{u}, \mathbf{v}) \geq F(\mathbf{u}) \\ Q(\mathbf{v}, \mathbf{v}) = F(\mathbf{v}) \end{cases} \quad (5)$$

具体地,惩罚函数 $\phi(u; a)$ 的受控优化算子可构建为:

$$q(u, v) = \frac{1}{2v} \phi'(v; a) u^2 + \frac{1}{2} v = \frac{1}{2v} \phi'(v; a) u^2 + C \quad (6)$$

式中: C 是相对于 v 的函数,与输入变量 u 无关。

2 编码器信号特征

编码器作为一种高精度机电设备,在旋转机械系统中广泛应用于角位移测量和运动控制。尤其在现代工业设备中,编码器不仅承担着控制反馈的重要角色,其输出信号还蕴含着反映设备运行状态的丰富信息。

尽管工业现场中存在多种类型的编码器,其输出信号的基本形式可统一^[20]为:

$$\varphi(t) = v_0 t + \sum_i g(t - iT_0) + \sum_m h_m(t) + w(t) \quad (7)$$

式中:第1项代表恒定速度 v_0 运行时的累计角位置;第2项为故障引起的周期性角度冲击特征,冲击特征周期为 T_0 。在本文中,利用一个周期的余弦信号来模拟故障引起的角度冲击特征,具体定义为:

$$g(t) = \frac{1}{2}A[\cos(2\pi ft) - 1] \quad (8)$$

式中:参数 A 和 f 决定了角度冲击特征的幅值和持续时间。

第3项为谐波成分,主要由啮合刚度波动、载荷不均、轴系不对中等因素诱发,其典型频率包括啮合频率、轴转频及其倍频等。第 m 阶谐波可建模为:

$$h_m = B_m \sin(2\pi f_m t + \varphi_m) \quad (9)$$

式中: B_m 与 φ_m 分别表示第 m 阶谐波的幅值和初始相位; f_m 为基频率。综合多个阶次谐波后,构成谐波干扰信号。

此外,最后一项是表示量化误差、传感器噪声或其他不确定扰动带来的随机干扰。

为实现旋转部件的有效故障识别,通常需将编码器记录的角位移信号转换为更能反映设备动态特性的运动学变量,其中最常用的是瞬时角速度。具体而言,可采用差分法对编码器信号进行处理,计算公式为:

$$IAS = \varphi'(t_k) = (\varphi(t_k + \Delta t) - \varphi(t_k)) / \Delta t \quad (10)$$

式中: $\Delta(t)$ 为采样时间间隔,等于信号采样频率的倒数; $\varphi'(t_k)$ 为在 t_k 时刻的瞬时角速度。

接下来,为进一步详细说明编码器信号的基本特性,构建了一个包含典型故障成分的仿真信号。表2列出了仿真编码器信号的参数。仿真编码器信号的时间长度为1 s,采样频率为5 000 Hz。图2显示了由机械故障引起的周期性故障冲击信号。

在上述设置基础上,将累计角位移、故障冲击、谐波成分以及背景噪声进行叠加,构造出仿真编码器信号,其时域波形如图3(a)所示。由于旋转速度设定为常数,编码器信号整体表现为近似线性增长的趋势,从时域上几乎无法直接识别出潜在故障特征。因此,有必要开展进一步的编码器信号分析,以从原始编码器信号中提取诊断信息,从而识别机械设备的健康状态。

表2 仿真编码器信号参数

Table 2 The parameters of the simulated encoder signal

信号成分	参数	数值
累计角位置	v_0 / rpm	600
故障	A / (°)	0.191
冲击	T_0 / s	0.1
	f / Hz	200
谐波成分	B_1, B_2, B_3, B_4, B_5 / (°)	0.95, 0.24, 0.57, 0.14, 0.04
	f_1, f_2, f_3, f_4, f_5 / Hz	10, 20, 30, 40, 50
	$\varphi_1, \varphi_2, \varphi_3, \varphi_4, \varphi_5$ / rad	$\pi/6, \pi/3, \pi/6, \pi/3, \pi/6$

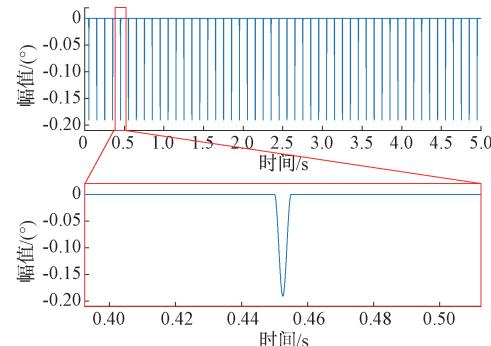


图2 仿真故障冲击

Fig. 2 Simulated fault impulses

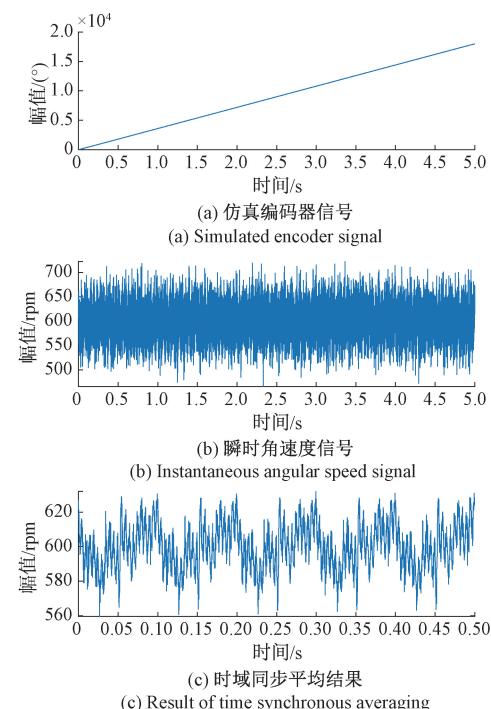


图3 仿真编码器信号及其预处理

Fig. 3 Simulated encoder signal and its preprocessing

首先,对原始编码器信号进行一阶差分处理,得到瞬时角速度信号,如图 3(b)所示。可以观察到,在含有多种干扰成分的条件下,时域内故障特征仍不明显,无法清晰辨识周期冲击信号。接着,使用时域同步平均方法以增强周期特征,处理结果如图 3(c)所示。尽管该方法利用了周期对齐平均的思想,有助于抑制随机噪声并强化周期成分,但由于故障冲击的能量相对较弱,且信号中仍混有强干扰成分,目标特征依然未能有效显现。

这一现象的根本原因在于:谐波干扰作为背景成分,往往具有比目标冲击更高的幅度,且频率可能为故障特征频率的整数倍,难以通过时域同步平均方法去除,从而掩盖冲击特征的真实响应。噪声与谐波的协同干扰,使得即使经过时域同步平均处理,非故障分量仍在信号中占据主导地位。因此,仅依赖传统时域平均难以实现对弱冲击特征的有效提取,亟需设计更具针对性和选择性的特征提取策略以提升诊断性能。

3 加权双域稀疏分解模型

3.1 模型构建

尽管时域同步平均方法在一定程度上能够增强信号中的周期性故障特征,但由于谐波干扰广泛存在,周期冲击信息常常被掩盖,导致微弱故障特征难以有效提取。为此,提出一种加权双域稀疏分解模型,通过联合利用时域与频域的稀疏性,实现周期冲击成分与谐波干扰的精确分离。

该模型的理论基础在于:故障冲击信号与谐波成分在时域、频域具有截然不同的结构特性。具体而言:

1) 周期冲击信号在时间域中呈现为周期性组稀疏结构,即仅有少量非零元素代表故障冲击,这些非零元素呈周期性分布,且往往成组出现;

2) 谐波成分在频率域中表现为少量集中的谱峰,具有明显的频域稀疏特性。

信号建模为 3 部分的叠加,即:

$$y = A_1 x_1 + x_2 + n \quad (11)$$

式中: y 表示原始信号; $A_1 x_1$ 表示谐波分量, A_1 表示逆傅里叶变换算子,用于将频域稀疏系数 x_1 映射回时域信号; x_2 表示故障引起的周期性冲击特征分量; n 表示背景噪声。

为实现周期性故障冲击特征与谐波成分的解耦,根据故障引起的周期性冲击特征与谐波成分在时域、频域具有不同形态特性的特点,构建加权双域稀疏分解模型,即:

$$\begin{aligned} \{x_1^*, x_2^*\} &= \operatorname{argmin}_{x_1, x_2} \{F(x_1, x_2) = \\ &\frac{1}{2} \|y - A_1 x_1 - x_2\|_2^2 + \lambda_1 \sum_n \phi([w_1]_n [x_1]_n; a_1) + \end{aligned}$$

$$\lambda_2 \sum_n \phi([w_2]_n \|b \odot [x_2]_{n,K}\|_2; a_2)\} \quad (12)$$

式中: λ_1 与 λ_2 为正则参数,分别控制频域与时间域的稀疏约束强度; \odot 表示元素点乘操作; 同时,将非凸惩罚函数 ϕ 设为 atan , 以最大程度地避免低估较大值,又能最大限度地促进系数的稀疏性。

频域稀疏正则项 $\sum_n \phi([w_1]_n [x_1]_n; a_1)$ 用于约束谐波成分在频域中的频域稀疏性, 周期组稀疏正则项 $\sum_n \phi([w_2]_n \|b \odot [x_2]_{n,K}\|_2; a_2)$ 用于约束故障冲击特征在时域中的周期性组稀疏性, 其综合了组稀疏结构与非凸惩罚函数, 有助于增强具有明确周期结构的冲击特征的识别能力。

其中, $[x_2]_{n,K}$ 表示第 n 个组, 该组共包含了 K 个元素, 定义为:

$$[x_2]_{n,K} = [[x_2]_n, [x_2]_{n+1}, \dots, [x_2]_{n+K-1}] \in \mathbf{R}^K \quad (13)$$

此外, b 是一个周期性二值向量, 用于约束冲击信号在时域中的周期性组稀疏特性, 其定义为:

$$b =$$

$$\underbrace{[1, \dots, 1, 0, \dots, 0, 1, \dots, 1, 0, \dots, 0, \dots, 1, \dots, 1, 0, \dots, 0]}_{K \text{ 个元素, } M \text{ 个故障周期}} \quad (14)$$

式中: N_1 表示单个故障周期中非零点的数量; N_0 为单个故障周期中零点的数量。此外, N_0, N_1 和 K 应该满足:

$$N_0 + N_1 = f_s / f_c \quad (15)$$

$$N_0 + N_1 = K / M \quad (16)$$

式中: f_s 是采样频率; f_c 表示故障特征频率; M 为周期二进制向量 b 中包含的故障周期个数。

此外,为进一步提升模型对稀疏先验的利用能力,受文献[21]提出的重加权 l_1 范数最小化方法的启发,在加权双域稀疏分解模型中引入了权重系数 w_1 和 w_2 。权重系数能够强化谐波成分及冲击特征在频域、时域中的稀疏先验及周期组稀疏先验结构,突出显著特征,从而使加权双域稀疏分解模型更具针对性地提取特征,提高两种特征的解耦分离精度,其表达式为:

$$[w_1]_n = \frac{1}{[x_1]_n + s} \quad (17)$$

$$[w_2]_n = \frac{1}{\|b \odot [x_2]_{n,K}\|_2 + s} \quad (18)$$

式中: $s > 0$ 是一个极小的正数, 用于避免分母为 0 的情况。

3.2 算法推导

在本节中,为了求解构建的无约束优化问题——加权双域稀疏分解模型,利用交替方法乘子法和受控极小化方法,推导了一种迭代求解算法。

首先,针对无约束优化问题(式(12)),引入满足 $\mathbf{u}_1 = \mathbf{x}_1$ 辅助变量 \mathbf{u}_1 ,并结合交替方向乘子法,可以得到迭代算法。通过不断迭代更新,能够获得优化问题(式(12))的最终解。

$$\begin{aligned} \mathbf{x}_1 &= \arg \min_{\mathbf{x}_1} \frac{1}{2} \|\mathbf{y} - \mathbf{A}_1 \mathbf{x}_1 - \mathbf{x}_2\|_2^2 + \\ \frac{\mu}{2} \|\mathbf{u}_1 - \mathbf{x}_1 - \mathbf{d}_1\|_2^2 \end{aligned} \quad (19)$$

$$\begin{aligned} \mathbf{x}_2 &= \arg \min_{\mathbf{x}_2} \frac{1}{2} \|\mathbf{y} - \mathbf{A}_1 \mathbf{x}_1 - \mathbf{x}_2\|_2^2 + \\ \lambda_2 \sum_n \phi([\mathbf{w}_2]_n \|\mathbf{b} \odot [\mathbf{x}_2]_{n,K}\|_2; a_2) \end{aligned} \quad (20)$$

$$\begin{aligned} \mathbf{u}_1 &= \arg \min_{\mathbf{u}_1} \frac{\mu}{2} \|\mathbf{u}_1 - \mathbf{x}_1 - \mathbf{d}_1\|_2^2 + \\ \lambda_1 \sum_n \phi([\mathbf{w}_1]_n [\mathbf{u}_1]_n; a_1) \end{aligned} \quad (21)$$

$$\mathbf{d}_1 = \mathbf{d}_1 - (\mathbf{u}_1 - \mathbf{x}_1) \quad (22)$$

式中: μ 为惩罚参数。

接下来,详细地说明优化子问题(式(19)~(21))的求解过程。

1)求解子问题(式(19))

优化子问题(式(19))是一个标准的最小二乘问题,其显式解为:

$$\mathbf{x}_1 = (\mathbf{A}_1^T \mathbf{A}_1 + \mu \mathbf{I})^{-1} (\mathbf{A}_1^T (\mathbf{y} - \mathbf{x}_2) + \mu (\mathbf{u}_1 - \mathbf{d}_1)) \quad (23)$$

根据矩阵逆引理,且逆傅里叶变换矩阵 \mathbf{A}_1 满足 $\mathbf{A}_1 \mathbf{A}_1^T = \mathbf{I}$,式(23)可以转换为:

$$\begin{aligned} \mathbf{x}_1 &= \frac{1}{\mu + 1} \mathbf{A}_1^T \left((\mathbf{y} - \mathbf{x}_2) - \frac{1}{\mu + 1} \mathbf{A}_1 (\mathbf{u}_1 - \mathbf{d}_1) \right) + \\ (\mathbf{u}_1 - \mathbf{d}_1) \end{aligned} \quad (24)$$

2)求解子问题(式(20))

为了最小化目标函数 $F(\mathbf{x}_2) = \frac{1}{2} \|\mathbf{y} - \mathbf{A}_1 \mathbf{x}_1 - \mathbf{x}_2\|_2^2 + \lambda_2 \sum_n \phi([\mathbf{w}_2]_n \|\mathbf{b} \odot [\mathbf{x}_2]_{n,K}\|_2; a_2)$,利用受控极小化方法推导出一种收敛算法。首先,目标函数 $F(\mathbf{x}_2)$ 的受控优化算子为:

$$\begin{aligned} Q(\mathbf{x}_2, \mathbf{v}_2) &= \frac{1}{2} \|\mathbf{y} - \mathbf{A}_1 \mathbf{x}_1 - \mathbf{x}_2\|_2^2 + \\ \lambda_2 \sum_n q([\mathbf{w}_2]_n \|\mathbf{b} \odot [\mathbf{x}_2]_{n,K}\|_2, [\mathbf{w}_2]_n \|\mathbf{b} \odot [\mathbf{v}_2]_{n,K}\|_2) = \\ \frac{1}{2} \|\mathbf{y} - \mathbf{A}_1 \mathbf{x}_1 - \mathbf{x}_2\|_2^2 + \frac{\lambda_2}{2} \sum_n [\boldsymbol{\gamma}_2(\mathbf{v}_2)]_{n,n} [\mathbf{x}_2]_n^2 + C \end{aligned} \quad (25)$$

式中: $\boldsymbol{\gamma}_2(\mathbf{v}_2)$ 为对角矩阵,定义为:

$$\begin{aligned} [\boldsymbol{\gamma}_2(\mathbf{v}_2)]_{n,n} &= \\ \sum_{k=1}^K [\mathbf{b}]_k \frac{\phi'([\mathbf{w}_2]_n \|\mathbf{b} \odot [\mathbf{v}_2]_{n-k+1,K}\|_2; a_2)}{[\mathbf{w}_2]_n \|\mathbf{b} \odot [\mathbf{v}_2]_{n-k+1,K}\|_2} [\mathbf{w}_2]_{n-k+1}^2 \end{aligned} \quad (26)$$

根据受控极小化方法的框架,最优解可以通过迭代式(27)获得,即:

$$\mathbf{x}_2^{(j+1)} = (\mathbf{y} - \mathbf{A}_1 \mathbf{x}_1^{(j)}) / (\lambda_2 \boldsymbol{\gamma}_2(\mathbf{x}_2^{(j)}) + 1) \quad (27)$$

3)求解子问题(式(21))

为了最小化目标函数 $F(\mathbf{u}_1) = \frac{\mu}{2} \|\mathbf{u}_1 - \mathbf{x}_1 - \mathbf{d}_1\|_2^2 +$

$\lambda_1 \sum_n \phi([\mathbf{w}_1]_n [\mathbf{u}_1]_n; a_1)$,利用受控极小化方法推导出一种收敛算法。首先,目标函数 $F(\mathbf{u}_1)$ 的受控优化算子为:

$$Q(\mathbf{u}_1, \mathbf{v}_1) = \frac{\mu}{2} \|\mathbf{u}_1 - \mathbf{x}_1 - \mathbf{d}_1\|_2^2 +$$

$$\lambda_1 \sum_n q([\mathbf{w}_1]_n [\mathbf{u}_1]_n, [\mathbf{w}_1]_n [\mathbf{v}_1]_n) = \frac{\mu}{2} \|\mathbf{u}_1 - \mathbf{x}_1 - \mathbf{d}_1\|_2^2 +$$

$$\frac{\lambda_1}{2} \sum_n [\boldsymbol{\gamma}_1(\mathbf{v}_1)]_{n,n} [\mathbf{u}_1]_n^2 + C \quad (28)$$

其中, $\boldsymbol{\gamma}_1(\mathbf{v}_1)$ 为对角矩阵,定义为:

$$[\boldsymbol{\gamma}_1(\mathbf{v}_1)]_{n,n} = \frac{\phi'([\mathbf{w}_1]_n [\mathbf{v}_1]_n; a_1)}{[\mathbf{w}_1]_n [\mathbf{v}_1]_n} [\mathbf{w}_1]_n^2 \quad (29)$$

根据受控极小化方法的框架,最优解可以通过迭代式(30)获得,即:

$$\mathbf{u}_1^{(j+1)} = (\mathbf{x}_1^{(j)} + \mathbf{d}_1^{(j)}) / \left(1 + \frac{\lambda_1 \boldsymbol{\gamma}_1(\mathbf{u}_1^{(j)})}{\mu} \right) \quad (30)$$

综上所述,算法1为求解稀疏优化问题(12)的迭代求解算法。利用提取的周期性冲击特征 \mathbf{x}_2 ,可以容易地识别潜在故障。

算法1:加权双域稀疏分解模型的迭代求解算法

- 输入: $\mathbf{y}, \mathbf{b}, iter, \mu, \lambda_1, \lambda_2$
- 对 $j \in [0, iter]$,执行循环
 - 利用式(24)计算 \mathbf{x}_1
 - 利用式(26)和(27)计算 \mathbf{x}_2
 - 利用式(29)和(30)计算 \mathbf{u}_1
 - 利用式(22)计算 \mathbf{d}_1
- 输出:故障冲击特征 \mathbf{x}_2 和谱波成分 $\mathbf{A}_1 \mathbf{x}_1$

3.3 基于加权双域稀疏分解模型的故障诊断流程

基于前文提出的加权双域稀疏分解模型,构建了一种基于编码器信号的编码器信号特征提取及故障诊断方法,其总体框架如图4所示。

该流程能够有效提取编码器信号中的周期性冲击特征,实现对机械系统运行状态的精准判断。具体步骤为:

1)测试原始编码器信号:从被监测设备中获取编码器信号,记录其采样频率及相关运行工况等参数,为后续处理提供数据支撑。

2)编码器信号预处理:利用一阶差分算法将原始编码器信号转换为瞬时角速度信号,以揭示运动过程中的

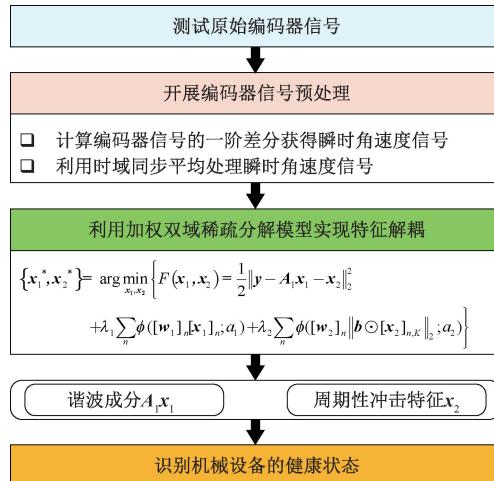


图 4 基于加权双域稀疏分解模型的故障诊断流程

Fig. 4 Flowchart of fault diagnosis based on the weighted bi-domain sparse decomposition model

微小扰动；随后，并对其进行时域同步平均处理，增强编码器信号中可能存在的周期性故障冲击特征。

3) 利用加权双域稀疏分解模型实现特征解耦：将预处理后的信号输入加权双域稀疏分解模型，利用其对时域中周期组稀疏性与频域中谐波稀疏性的联合建模能力，实现周期性故障冲击特征与谐波干扰成分的有效解耦与提取。

4) 识别机械设备的健康状态：利用提取出的周期性故障冲击特征，判断机械设备的健康状态。

4 仿真分析

为验证本文所提出的加权双域稀疏分解模型在编码器信号故障特征提取中的有效性，本章将其应用于第 2 章所构建的仿真编码器信号，并与两种主流信号处理方法——调 Q 因子小波变换 (tunable Q wavelet transform, TQWT) 分解方法和最大相关峭度解卷积 (maximum correlated kurtosis deconvolution, MCKD) 方法进行对比分析。

首先，为进一步增强周期性冲击信号的可观测性，将加权双域稀疏分解模型方法应用于图 3(c) 中经预处理的瞬时角速度信号。处理结果如图 5 所示，其中图 5(a) 为分离得到的谐波成分，图 5(b) 为提取出的故障冲击特征。可以明显观察到，提出方法成功实现了周期性冲击成分与谐波干扰成分的清晰解耦。特别地，图 5(b) 中以 0.1 s 为周期出现的冲击特征，清晰揭示了故障的周期性结构。这种良好的分离效果得益于加权双域稀疏分解模型在构建过程中对冲击信号的周期组稀疏性和谐波信号的频域稀疏性的联合建模，有效提升了故障特征提取的精度与鲁棒性。

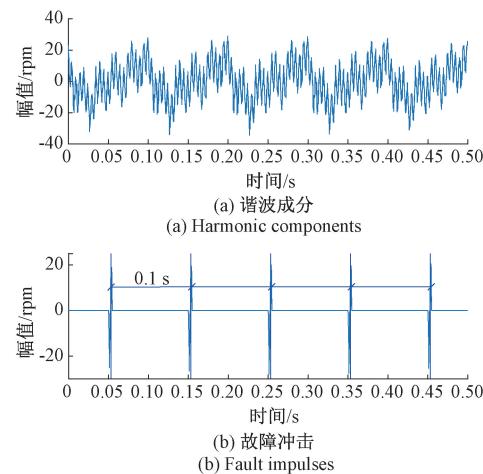


图 5 提出方法分析结果

Fig. 5 Result of the proposed method

为了验证所提方法的优势，图 6 展示了采用 TQWT 分解方法的处理结果。图 6(a) 与图 6(b) 分别对应谐波成分与冲击分量的提取结果。尽管该方法在一定程度上分离了信号的不同成分，但故障冲击特征仍受到明显的谐波成分与随机干扰的影响，导致冲击信号的清晰度与可识别性明显下降，影响故障的可靠诊断。进一步地，图 7 展示了 MCKD 方法的分析结果。尽管该方法旨在通过增强信号中的尖峰结构以识别潜在冲击特征，但从图 7 中可以看出，其提取的脉冲序列存在幅值失真，且仍受到显著噪声的干扰。这一结果表明，MCKD 方法在面对编码器信号这类高干扰、成分复杂的信号时，表现出一定的局限性，难以实现对周期性冲击特征的稳定增强与噪声的有效抑制，进而限制了其在故障特征提取中的实用性和鲁棒性。

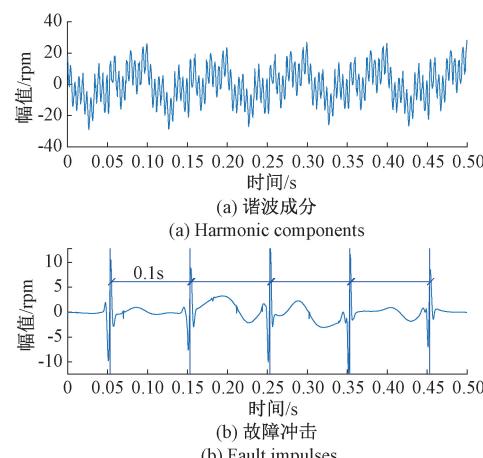


图 6 TQWT 分解方法分析结果

Fig. 6 Results of the TQWT-based decomposition method

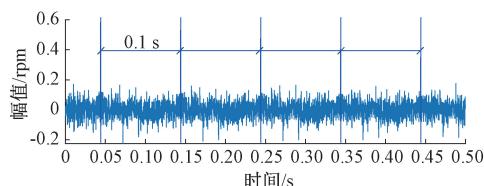


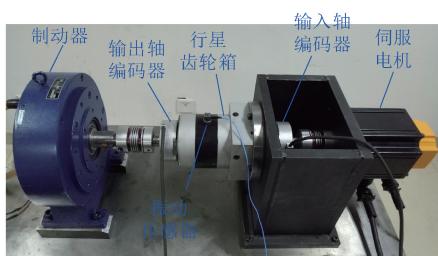
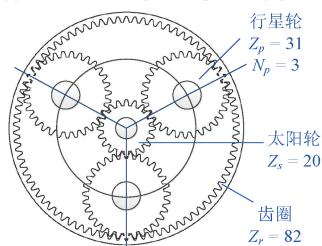
图 7 最大相关峭度解卷积分析结果

Fig. 7 Results of maximum correlated kurtosis deconvolution

5 实验验证

为验证所提出的加权双域稀疏分解模型在实际工况下的故障诊断性能和鲁棒性,本章将其应用于行星齿轮箱实验台的编码器信号分析,并与 TQWT 分解方法及 MCKD 方法进行对比分析。

图 8(a)展示了实验平台的整体结构,由伺服电机、行星齿轮箱、联轴器及制动器等组成。该平台可调节转速与载荷,模拟实际工程中行星齿轮箱在不同工况下的运行状态。编码器分别安装于输入轴与输出轴,用于采集旋转位置信息。图 8(b)为齿轮箱结构示意图,其内部结构包括 1 个太阳轮(齿数为 20)、3 个等间隔布置的行星轮(齿数为 31)及 1 个齿圈(齿数为 82)。太阳轮与输入轴连接,带动行星轮旋转,行星轮通过支架与输出轴连接实现功率传递,齿圈固定不动。此外,采用海德汉 IK220 高性能计数卡进行数据采集,采样频率为 5 000 Hz。

(a) 行星齿轮箱实验台图片
(a) Picture of test rig of the planetary gearbox(b) 行星齿轮箱结构示意图
(b) Diagram of the planetary gearbox图 8 行星齿轮箱实验台
Fig. 8 Test rig of the planetary gearbox

实验设置伺服电机转速为 20 Hz,根据行星齿轮箱结构参数可以计算太阳轮、行星轮及齿圈的故障特征频率,如表 3 所示,作为后续故障定位依据。具体地, f_i 和 f_o 分别为伺服电机和输出轴转速, f_p , f_s 和 f_r 分别为太阳轮、行星轮和齿圈的故障特征频率。

表 3 行星齿轮箱故障特征频率

Table 3 Fault characteristic frequencies of

planetary gearbox					(Hz)
f_i	f_o	f_p	f_s	f_r	
20	3.921 6	10.373 2	48.235 3	11.764 7	

在实验平台上测试了单齿磨损故障的行星轮,如图 9 所示,并采集了输入轴和输出轴的编码器信号,以评估提出方法的故障诊断性能。值得说明的是,输入轴和输出轴编码器信号均能够用于行星齿轮箱的故障诊断,本章仅以输出轴编码器信号分析为例说明提出方法的有效性。



图 9 单齿磨损行星轮

Fig. 9 Planet gear with single tooth wear

图 10(a)展示了来自输出轴编码器的原始信号,持续时间为 3 s。由于编码器输出为累计角位移,其波形呈现出近似线性上升趋势,无法直接揭示潜在的故障信息。为增强信号中潜藏的瞬态特征,对编码器信号进行预处理操作。首先,对原始编码器信号进行一阶差分,获取瞬时角速度信号,结果如图 10(b)所示。然而,由于信号中含有大量谐波成分和背景噪声,周期性故障冲击特征仍被严重淹没,难以识别。这表明单纯通过差分操作得到的瞬时角速度未能有效揭示与齿轮故障相关的诊断信息,其原因主要在于差分操作虽可增强故障特征,但同时也扩大了噪声与谐波的干扰效应。为了进一步提高周期性冲击特征的可识别性,对瞬时角速度信号应用时域同步平均处理,以抑制非周期成分、增强目标故障频率的表达。图 10(c)为时域同步平均处理结果,其中周期冲击特征仍受到谐波及其倍频干扰的显著掩盖,无法直接用于故障定位与识别。

基于上述信号预处理,进一步将图 10(c)中的信号输入到本文提出的加权双域稀疏分解模型中,以实现周

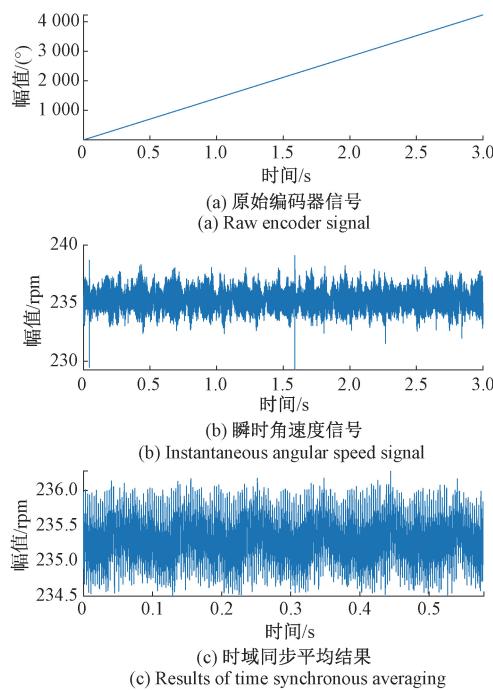


图 10 原始编码器信号及其预处理

Fig. 10 Raw encoder signal and its preprocessing

期性冲击特征与谐波成分的解耦分离。图 11 展示了提出方法的分解结果, 可以看出谐波成分及周期性故障冲击特征被准确分离。特别地, 图 11(b)成功提取出具有明显周期结构的冲击特征信号。值得注意的是, 提取出的冲击信号的时间间隔 $1/f_p$ 与行星轮故障周期一致, 表明准确识别了行星轮故障, 验证了该方法在故障识别精度和信噪比增强方面的显著优势。

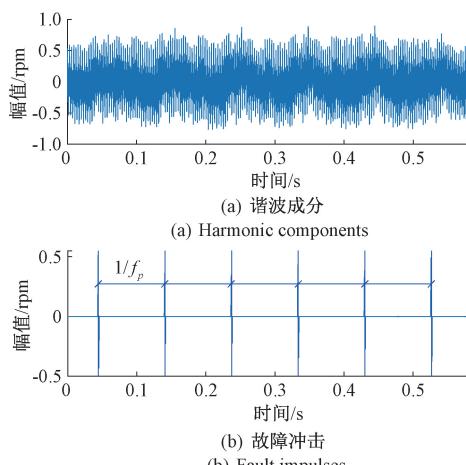


图 11 提出方法分析结果

Fig. 11 Results of the proposed method

为进一步评估提出方法的性能, 引入 TQWT 分解方法与 MCKD 方法作为对比手段。图 12 所示为 TQWT 分

解方法的结果, 虽然其也实现了信号的分离, 但图 12(b)中的冲击特征仍受到明显随机噪声干扰, 仅部分周期性特征得以显现。这表明 TQWT 分解方法在提取微弱周期性冲击方面存在一定局限性, 主要由于其未能显式约束冲击信号的周期组稀疏结构, 导致谐波成分与瞬态冲击成分的时域混叠未被充分抑制。

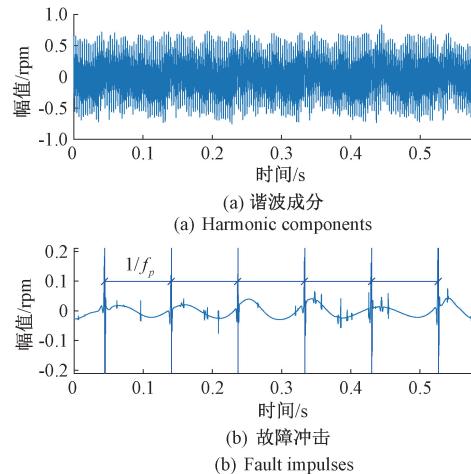


图 12 TQWT 分解方法分析结果

Fig. 12 Results of TQWT-based decomposition method

进一步地, 图 13 呈现了 MCKD 方法的分析结果。除了故障引起的周期性冲击特征外, 还可以观察到其他干扰冲击特征的存在, 且背景噪声复杂并存在幅值失真, 降低了故障冲击特征的可识别性, 限制其在弱故障识别中的适用性。

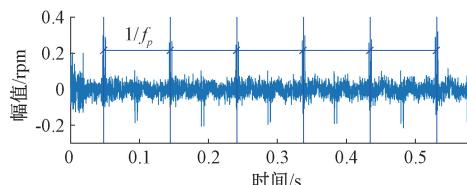


图 13 最大相关峭度解卷积分析结果

Fig. 13 Results of maximum correlated kurtosis deconvolution

综上所述, 本文提出的加权双域稀疏分解方法在强噪声背景下依然展现出良好的周期冲击识别能力, 显著优于 TQWT 分解方法与 MCKD 方法, 验证了其在编码器信号故障诊断中的有效性与适用性。

6 结 论

针对编码器信号中故障冲击特征微弱且易被多种干扰成分叠加的问题, 提出了一种基于加权双域稀疏分解模型的特征提取与行星齿轮箱故障诊断方法, 实现了在复杂噪声环境下的高效特征提取与故障识别。主要结论为:

1) 通过深入分析故障冲击特征与干扰成分在时域和频域的形态差异,结合权重系数、非凸惩罚函数与二进制周期向量,构建了分别约束周期性组稀疏性和频域稀疏性的非凸正则项,促进周期性故障冲击特征与谐波干扰成分的高精度分离,有效提升了信号的特征辨识能力。

2) 基于交替方向乘子法与受控极小化方法,设计了一种迭代求解算法,实现了加权双域稀疏分解模型的高效求解。

3) 通过分析仿真信号及行星齿轮箱实验台信号,表明所提方法在故障冲击特征提取准确性方面优于传统方法,能够有效识别行星齿轮箱故障,展现出良好的工程应用潜力。

参考文献

- [1] 江星星,宋秋昱,杜贵府,等.变分模式分解方法研究与应用综述[J].仪器仪表学报,2023,44(1):55-73.
- JIANG X X, SONG Q Y, DU G F, et al. Review on research and application of variational mode decomposition[J]. Chinese Journal of Scientific Instrument, 2023, 44(1): 55-73.
- [2] 石明宽,丁传仓,王锐,等.面向变工况下工业流数据故障诊断的持续迁移学习系统[J].仪器仪表学报,2024,45(4):10-16.
- SHI M K, DING CH C, WANG R, et al. Continuous transfer learning system for fault diagnosis of industrial stream data under variable operating conditions [J]. Chinese Journal of Scientific Instrument, 2024, 45(4): 10-16.
- [3] 张锐,刘婷婷,王燕,等.基于FBSE-ESEWT的齿轮故障诊断方法[J].电子测量与仪器学报,2025,39(4):234-246.
- ZHANG R, LIU T T, WANG Y, et al. Gear fault diagnosis method based on FBSE-ESEWT[J]. Journal of Electronic Measurement and Instrumentation, 2025, 39(4): 234-246.
- [4] 万佳诚,曾宪文,李靖超.基于参数优化VMD与宽卷积神经网络的齿轮箱故障诊断[J].电子测量技术,2025,48(10):25-32.
- WAN J CH, ZENG X W, LI J CH. Fault diagnosis for gearbox based on parameter optimization VMD combined with wide convolutional neural network [J]. Electronic Measurement Technology, 2025, 48(10): 25-32.
- [5] YANG X M, GUO Y, WANG H W. Encoder signal

analysis and its application in gear fault detection [J]. IEEE Transactions on Instrumentation and Measurement, 2023, 73: 1-9.

- [6] MIAO Y H, ZHANG B Y, YI Y G, et al. Application of improved reweighted singular value decomposition for gearbox fault diagnosis based on built-in encoder information[J]. Measurement, 2021, 168: 108295.
- [7] OU SH D, ZHAO M, LI S, et al. Online shock sensing for rotary machinery using encoder signal[J]. Mechanical Systems and Signal Processing, 2023, 182: 109559.
- [8] FENG ZH P, GAO AO R, LI K Q, et al. Planetary gearbox fault diagnosis via rotary encoder signal analysis[J]. Mechanical Systems and Signal Processing, 2021, 149: 107325.
- [9] 刘振,孔德同,赵明,等.裂纹损伤行星轮系的编码器信号响应特性[J].西安交通大学学报,2020,54(2):43-50.
- LIU ZH, KONG D T, ZHAO M, et al. Encoder signal response characteristics of planetary gear sets with cracks[J]. Journal of Xi'an Jiaotong University, 2020, 54(2): 43-50.
- [10] LI B, ZHANG X N, WU J L. New procedure for gear fault detection and diagnosis using instantaneous angular speed[J]. Mechanical Systems and Signal Processing, 2017, 85: 415-428.
- [11] ZHAO Z R, WANG X L, HU Y Q, et al. Grease-lubricated triboelectric instantaneous angular speed sensor integrated with signal processing circuit for bearing fault diagnosis[J]. Nano Energy, 2023, 117: 108871.
- [12] DING CH C, ZHAO M, LIN J, et al. Sparsity-based algorithm for condition assessment of rotating machinery using internal encoder data [J]. IEEE Transactions on Industrial Electronics, 2019, 67(9): 7982-7993.
- [13] WANG B X, DING CH C. Transient feature identification from internal encoder signal for fault detection of planetary gearboxes under variable speed conditions[J]. Measurement, 2021, 171: 108761.
- [14] 欧曙东,赵明,周涛,等.基于编码器信号的低转速行星齿轮箱故障诊断技术[J].中国电机工程学报,2021,41(5):1885-1894.
- OU SH D, ZHAO M, ZHOU T, et al. Fault diagnosis technology for low-speed planetary gearbox based on

- encoder signals [J]. Proceedings of the CSEE, 2021, 41(5): 1885-1894.
- [15] 田田, 郭瑜, 樊家伟, 等. 基于编码器信号自适应MOMEDA的太阳轮故障检测[J]. 振动、测试与诊断, 2024, 44(6): 1173-1180, 1249.
- TIAN T, GUO Y, FAN J W, et al. Sun gear fault detection via adaptive MOMEDA based on encoder signal[J]. Journal of Vibration, Measurement & Diagnosis, 2024, 44(6): 1173-1180, 1249.
- [16] 张佳鑫, 郭瑜. 基于ASA-MCKD的IAS信号针齿故障特征提取研究[J]. 机电工程, 2025, 42(6): 1110-1117.
- ZHANG J X, GUO Y. Pin tooth fault feature extraction of IAS signal based on ASA-MCKD [J]. Journal of Mechanical & Electrical Engineering, 2025, 42(6): 1110-1117.
- [17] ZUO J Y, MIAO Y H, ZHANG B Y, et al. Cyclostationary feature mode decomposition and its application in fault diagnosis of planetary gearboxes via built-in information [J]. IEEE Sensors Journal, 2023, 24(2): 1129-1139.
- [18] ZHANG Z W, HUANG W G, WANG J, et al. Morphological component analysis under non-convex smoothing penalty framework for gearbox fault diagnosis[J]. ISA Transactions, 2023, 143: 525-535.
- [19] 王爽, 丁传仓, 曹懿, 等. 加权多尺度卷积稀疏表示及其在滚动轴承复合故障诊断中的应用[J]. 仪器仪表学报, 2024, 45(5): 197-207.
- WANG SH, DING CH C, CAO Y, et al. Weighted multiscale convolutional sparse representation and its application in rolling bearing compound fault diagnosis[J]. Chinese Journal of Scientific Instrument, 2024, 45(5): 197-207.
- [20] OU SH D, ZHAO M, ZHOU T, et al. An encoder signal-based approach for low-speed planetary gearbox fault diagnosis[J]. Measurement Science and Technology, 2021, 32(5): 054005.
- [21] CANDES E J, WAKIN M B, BOYD S P. Enhancing sparsity by reweighted ℓ_1 minimization [J]. Journal of Fourier Analysis and Applications, 2008, 14: 877-905.

作者简介



张建峰,2000 年于扬州大学获得学士学位,2012 年于江苏大学获得硕士学位,现为南通理工学院副教授,主要研究方向为汽车零部件性能测试、智能故障诊断和剩余寿命预测。

E-mail:zhangmajinjin@sohu.com

Zhang Jianfeng received his B. Sc. degree from Yangzhou University in 2000, and his M. Sc. degree from Jiangsu University in 2012. He is currently an associate professor at Nantong Institute of Technology. His main research interests include performance testing, intelligent fault diagnosis, and remaining useful life prediction of automotive components.



丁传仓(通信作者),2016 年于山东大学获得学士学位,2021 年于西安交通大学获得博士学位,现为苏州大学副教授,主要研究方向为机械信号分析与设备状态监测。

E-mail:ecding@suda.edu.cn

Ding Chuancang (Corresponding author) received his B. Sc. degree from Shandong University in 2016, and his Ph. D. degree from Xi'an Jiaotong University in 2021. He is currently an associate professor at Soochow University. His main research interests include mechanical signal analysis and machine condition monitoring.